SEQUENCE LISTING

<110> University of Rochester
      Maquat, Lynne E.

<120> NONSENSE-MEDIATED MRNA DECAY

<130> 21108.0023P1

<150> 60/405,602
<151> 2002-08-22

<160> 38

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 1
gcctattggt ctattttccc   20

<210> 2
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 2
cctgaagttc tcaggatc   18

<210> 3
<211> 32

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 3
atctggcacc acaccttcta caatgagctg cg   32

<210> 4
<211> 32
<212> DNA
<213> Artificial Sequence

```
<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 4
cgtcatactc ctgcttgctg atccacatct gc                                  32


<210> 5
<211> 19
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 5
tgcaaggagt ttcatcctg                                                 19


<210> 6
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 6
agaatcagta gtttaacaca c                                              21


<210> 7
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 7
tgagcatagt tattaatagc ag                                             22


<210> 8
<211> 77
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 8
gctagctcga gaccggtgcc accatggact acaaagacga tgacgacaag gcggaagggc    60
tggagcgtgt gcggatc                                                   77


<210> 9
<211> 43
<212> DNA
<213> Artificial Sequence
```

```
<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 9
tttaaacccg gcctgcgggg ccagagtagc caggatcccg cgc          43


<210> 10
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 10
tgaccttcag cgcctcgg                                       18


<210> 11
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 11
ctccgagtcc ctctgcc                                        17


<210> 12
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 12
ggcaaaggct ctgagaagc                                      19


<210> 13
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 13
ccgaggtccc aaaggcg                                        17


<210> 14
<211> 39
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 14
atcgaagatc tggatccaag gtcgggcagg agagggcct   39

<210> 15
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 15
tacacaaagc aatgtccatt acatgccacg gtgtttcgtc ctttccacaa gatatataa   59

<210> 16
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 16
cgaaatctag aaaaaagtgg catgtaatgg acattgccta cacaaagc   48

<210> 17

<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 17
gcugcagcag aacaggccat t   21

<210> 18
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 18
guacaaccca ggauaugugt t   21

<210> 19
<211> 59
<212> DNA
<213> Artificial Sequence

```
<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 19
tacacaaaca gggctgttct tcgagatgcg gtgtttcgtc ctttccacaa gatatataa        59


<210> 20
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 20
cgaaatctag aaaaaagcat ctcgaagaac agccctgcta cacaaaca                    48


<210> 21
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 21
tacacaaagc aatgtccgtt gcatgccacg gtgtttcgtc ctttccacaa gatatataa        59


<210> 22
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 22
cgaaatctag aaaaaagtgg catgcaacgg acattgccta cacaaagc                    48


<210> 23
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 23
tacacaaagt tcagaggctg tgtcataacg gtgtttcgtc ctttccacaa gatatataa        59


<210> 24
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
```

<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 24
cgaaatctag aaaaaagtta tgacacagcc tctgaaccta cacaaagt   48

<210> 25
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 25
tacacaaacc aaggcacttg ttggcagtcg gtgtttcgtc ctttccacaa gatatataa   59

<210> 26
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 26
cgaaatctag aaaaaagact gccaacaagt gccttggcta cacaaacc   48

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 27
gcagcgagca actgagaagc   20

<210> 28
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 28
gggtttagtg gtacttgtga gc   22

<210> 29
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =

synthetic construct

<400> 29
gactgagccg atcccgcgc   19

<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 30
gcagtaacgg cagacttctc   20

<210> 31
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     'synthetic construct

<400> 31
cctttcctgc tcttgcctg   19

<210> 32
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 32
gctttttatt tgtcagaaga cag   23

<210> 33
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 33
atctggcacc acaccttcta caatgagctg   30

<210> 34
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
     synthetic construct

<400> 34
cgtcatactc ctgcttgctg atccacatct   30

<210> 35
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 35
atgacttcga aagtttat   18

<210> 36
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 36
ttcagatttg atcaacgca   19

<210> 37
<211> 1419
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 37
Met Ala Glu Gly Leu Glu Arg Val Arg Ile Ser Ala Ser Glu Leu Arg
 1   5   10   15
Gly Ile Leu Ala Thr Leu Ala Pro Gln Ala Gly Ser Arg Glu Asn Met
            20   25   30
Lys Glu Leu Lys Glu Ala Arg Pro Arg Lys Asp Asn Arg Arg Pro Asp
        35   40   45
Leu Glu Ile Tyr Lys Pro Gly Leu Ser Arg Leu Arg Asn Lys Pro Lys
    50   55   60
Ile Lys Glu Pro Pro Gly Ser Glu Glu Phe Lys Asp Glu Ile Val Asn
65   70   75   80
Asp Arg Asp Cys Ser Ala Val Glu Asn Gly Thr Gln Pro Val Lys Asp
                85   90   95
Val Cys Lys Glu Leu Asn Asn Gln Glu Gln Asn Gly Pro Ile Asp Pro
            100   105   110
Glu Asn Asn Arg Gly Gln Glu Ser Phe Pro Arg Thr Ala Gly Gln Glu
            115   120   125
Asp Arg Ser Leu Lys Ile Ile Lys Arg Thr Lys Lys Pro Asp Leu Gln
    130   135   140
Ile Tyr Gln Pro Gly Arg Arg Leu Gln Thr Val Ser Lys Glu Ser Ala
145   150   155   160
Ser Arg Val Glu Glu Glu Glu Val Leu Asn Gln Val Glu Gln Leu Arg
                165   170   175

```
Val Glu Glu Asp Glu Cys Arg Gly Asn Val Ala Lys Glu Glu Val Ala
        180                 185             190
Asn Lys Pro Asp Arg Ala Glu Ile Glu Lys Ser Pro Gly Gly Gly Arg
        195                 200             205
Val Gly Ala Ala Lys Gly Glu Lys Gly Lys Arg Met Gly Lys Gly Glu
        210                 215             220
Gly Val Arg Glu Thr His Asp Asp Pro Ala Arg Gly Arg Pro Gly Ser
225                 230             235                 240
Ala Lys Arg Tyr Ser Arg Ser Asp Lys Arg Arg Asn Arg Tyr Arg Thr
                245             250                 255
Arg Ser Thr Ser Ser Ala Gly Ser Asn Asn Ser Ala Glu Gly Ala Gly
        260                 265             270
Leu Thr Asp Asn Gly Cys Arg Arg Arg Arg Gln Asp Arg Thr Lys Glu
        275                 280             285
Arg Pro Pro Leu Lys Lys Gln Val Ser Val Ser Ser Thr Asp Ser Leu
        290                 295             300
Asp Glu Asp Arg Ile Asp Glu Pro Asp Gly Leu Gly Pro Arg Arg Ser
305                 310             315                 320
Ser Glu Arg Lys Arg His Leu Glu Arg Asn Trp Ser Gly Arg Gly Glu
                325             330                 335
Gly Glu Gln Lys Thr Ser Ala Lys Glu Tyr Arg Gly Thr Leu Arg Val
        340                 345             350
Thr Phe Asp Ala Glu Ala Met Asn Lys Glu Ser Pro Met Val Arg Ser
        355                 360             365
Ala Arg Asp Asp Met Asp Arg Gly Lys Pro Asp Lys Gly Leu Ser Ser
        370                 375             380
Gly Gly Lys Gly Ser Glu Lys Gln Glu Ser Lys Asn Pro Lys Gln Glu
385                 390             395                 400
Leu Arg Gly Arg Gly Arg Gly Ile Leu Ile Leu Pro Ala His Thr Thr
                405             410             415
Leu Ser Val Asn Ser Ala Gly Ser Pro Glu Ser Ala Pro Leu Gly Pro
        420                 425             430
Arg Leu Leu Phe Gly Ser Gly Ser Lys Gly Ser Arg Ser Trp Gly Arg
        435                 440             445
Gly Gly Thr Thr Arg Arg Leu Trp Asp Pro Asn Asn Pro Asp Gln Lys
        450                 455             460
Pro Ala Leu Lys Thr Gln Thr Pro Gln Leu His Phe Leu Asp Thr Asp
465                 470             475                 480
Asp Glu Val Ser Pro Thr Ser Trp Gly Asp Ser Arg Gln Ala Gln Ala
                485             490                 495
Ser Tyr Tyr Lys Phe Gln Asn Ser Asp Asn Pro Tyr Tyr Tyr Pro Arg
                500             505             510
Thr Pro Gly Pro Ala Ser Gln Tyr Pro Tyr Thr Gly Tyr Asn Pro Leu
        515                 520             525
Gln Tyr Pro Val Gly Pro Thr Asn Gly Val Tyr Pro Gly Pro Tyr Tyr
        530                 535             540
Pro Gly Tyr Pro Thr Pro Ser Gly Gln Tyr Val Cys Ser Pro Leu Pro
545                 550             555                 560
Thr Ser Thr Met Ser Pro Glu Glu Val Glu Gln His Met Arg Asn Leu
                565             570             575
Gln Gln Gln Glu Leu His Arg Leu Leu Arg Val Ala Asp Asn Gln Glu
        580             585             590
Leu Gln Leu Ser Asn Leu Leu Ser Arg Asp Arg Ile Ser Pro Glu Gly
        595                 600             605
Leu Glu Lys Met Ala Gln Leu Arg Ala Glu Leu Leu Gln Leu Tyr Glu
        610             615             620
Arg Cys Ile Leu Leu Asp Ile Glu Phe Ser Asp Asn Gln Asn Val Asp
625                 630             635                 640
Gln Ile Leu Trp Lys Asn Ala Phe Tyr Gln Val Ile Glu Lys Phe Arg
                645             650             655
```

9

```
Gln Leu Val Lys Asp Pro Asn Val Glu Asn Pro Glu Gln Ile Arg Asn
              660                   665                   670
Arg Leu Leu Glu Leu Leu Asp Glu Gly Ser Asp Phe Phe Asp Ser Leu
              675                   680                   685
Leu Gln Lys Leu Gln Val Thr Tyr Lys Phe Lys Leu Glu Asp Tyr Met
              690                   695                   700
Asp Gly Leu Ala Ile Arg Ser Lys Pro Leu Arg Lys Thr Val Lys Tyr
705                   710                   715                   720
Ala Leu Ile Ser Ala Gln Arg Cys Met Ile Cys Gln Gly Asp Ile Ala
                  725                   730                   735
Arg Tyr Arg Glu Gln Ala Ser Asp Thr Ala Asn Tyr Gly Lys Ala Arg
              740                   745                   750
Ser Trp Tyr Leu Lys Ala Gln His Ile Ala Pro Lys Asn Gly Arg Pro
              755                   760                   765
Tyr Asn Gln Leu Ala Leu Leu Ala Val Tyr Thr Arg Arg Lys Leu Asp
              770                   775                   780
Ala Val Tyr Tyr Tyr Met Arg Ser Leu Ala Ala Ser Asn Pro Ile Leu
785                   790                   795                   800
Thr Ala Lys Glu Ser Leu Met Ser Leu Phe Glu Glu Thr Lys Arg Lys
                  805                   810                   815
Ala Glu Gln Met Glu Lys Lys Gln His Glu Glu Phe Asp Leu Ser Pro
              820                   825                   830
Asp Gln Trp Arg Lys Gly Lys Lys Ser Thr Phe Arg His Val Gly Asp
              835                   840                   845
Asp Thr Thr Arg Leu Glu Ile Trp Ile His Pro Ser His Pro Arg Ser
850                   855                   860
Ser Gln Gly Thr Glu Ser Gly Lys Asp Ser Glu Gln Glu Asn Gly Leu
865                   870                   875                   880
Gly Ser Leu Ser Pro Ser Asp Leu Asn Lys Arg Phe Ile Leu Ser Phe
                  885                   890                   895
Leu His Ala His Gly Lys Leu Phe Thr Arg Ile Gly Met Glu Thr Phe
              900                   905                   910
Pro Ala Val Ala Glu Lys Val Leu Lys Glu Phe Gln Val Leu Leu Gln
              915                   920                   925
His Ser Pro Ser Pro Ile Gly Ser Thr Arg Met Leu Gln Leu Met Thr
              930                   935                   940
Ile Asn Met Phe Ala Val His Asn Ser Gln Leu Lys Asp Cys Phe Ser
945                   950                   955                   960
Glu Glu Cys Arg Ser Val Ile Gln Glu Gln Ala Ala Ala Leu Gly Leu
                  965                   970                   975
Ala Met Phe Ser Leu Leu Val Arg Arg Cys Thr Cys Leu Leu Lys Glu
                  980                   985                   990
Ser Ala Lys Ala Gln Leu Ser Ser Pro Glu Asp Gln Asp Asp Gln Asp
              995                   1000                  1005
Asp Ile Lys Val Ser Ser Phe Val Pro Asp Leu Lys Glu Leu Leu Pro
         1010                  1015                  1020
Ser Val Lys Val Trp Ser Asp Trp Met Leu Gly Tyr Pro Asp Thr Trp
1025                  1030                  1035                  1040
Asn Pro Pro Pro Thr Ser Leu Asp Leu Pro Ser His Val Ala Val Asp
                  1045                  1050                  1055
Val Trp Ser Thr Leu Ala Asp Phe Cys Asn Ile Leu Thr Ala Val Asn
                  1060                  1065                  1070
Gln Ser Glu Val Pro Leu Tyr Lys Asp Pro Asp Asp Leu Thr Leu
              1075                  1080                  1085
Leu Ile Leu Glu Glu Asp Arg Leu Leu Ser Gly Phe Val Pro Leu Leu
         1090                  1095                  1100
Ala Ala Pro Gln Asp Pro Cys Tyr Val Glu Lys Thr Ser Asp Lys Val
1105                  1110                  1115                  1120
Ile Ala Ala Asp Cys Lys Arg Val Thr Val Leu Lys Tyr Phe Leu Glu
                  1125                  1130                  1135
```

10

```
Ala Leu Cys Gly Gln Glu Glu Pro Leu Leu Ala Phe Lys Gly Gly Lys
          1140              1145              1150
Tyr Val Ser Val Ala Pro Val Pro Asp Thr Met Gly Lys Glu Met Gly
          1155              1160              1165
Ser Gln Glu Gly Thr Arg Leu Glu Asp Glu Glu Glu Asp Val Val Ile
          1170              1175              1180
Glu Asp Phe Glu Glu Asp Ser Glu Ala Glu Gly Ser Gly Gly Glu Asp
1185              1190              1195              1200
Asp Ile Arg Glu Leu Arg Ala Lys Lys Leu Ala Leu Ala Arg Lys Ile
          1205              1210              1215
Ala Glu Gln Gln Arg Arg Gln Glu Lys Ile Gln Ala Val Leu Glu Asp
          1220              1225              1230
His Ser Gln Met Arg Gln Met Glu Leu Glu Ile Arg Pro Leu Phe Leu
          1235              1240              1245
Val Pro Asp Thr Asn Gly Phe Ile Asp His Leu Ala Ser Leu Ala Arg
          1250              1255              1260
Leu Leu Glu Ser Arg Lys Tyr Ile Leu Val Val Pro Leu Ile Val Ile
1265              1270              1275              1280
Asn Glu Leu Asp Gly Leu Ala Lys Gly Gln Glu Thr Asp His Arg Ala
          1285              1290              1295
Gly Gly Tyr Ala Arg Val Val Gln Glu Lys Ala Arg Lys Ser Ile Glu
          1300              1305              1310
Phe Leu Glu Gln Arg Phe Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu
          1315              1320              1325
Thr Ser Arg Gly Asn Glu Leu Glu Ser Ile Ala Phe Arg Ser Glu Asp
          1330              1335              1340
Ile Thr Gly Gln Leu Gly Asn Asn Asp Asp Leu Ile Leu Ser Cys Cys
1345              1350              1355              1360
Leu His Tyr Cys Lys Asp Lys Ala Lys Asp Phe Met Pro Ala Ser Lys
          1365              1370              1375
Glu Glu Pro Ile Arg Leu Leu Arg Glu Val Val Leu Leu Thr Asp Asp
          1380              1385              1390
Arg Asn Leu Arg Val Lys Ala Leu Thr Arg Asn Val Pro Val Arg Asp
          1395              1400              1405
Ile Pro Ala Phe Leu Thr Trp Ala Gln Val Gly
          1410              1415
```

<210> 38
<211> 5965
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

<400> 38
```
cctggctgcg cgcggcggtg gcggagccgc tacggctgta gcagcagccg cgaagatggc     60
ggaagggctg gagcgtgtgc ggatctccgc gtcggagctg cgcgggatcc tggctactct    120
ggccccgcag gccgggagca gagaaacat gaaggaatta aggaggcca ggccgcgcaa      180
agataacagg cgtccagatc tggaaatcta taagcctggc ctttctcggc taaggaacaa    240
gcccaaaatc aaggaacccc ctgggagtga ggaattcaaa gatgaaattg ttaatgaccg    300
agattgctct gctgttgaaa atggtacaca gcccgttaaa gatgtctgca aggaactgaa    360
caaccaagag cagaatggtc ctatagaccc agaaaataat cggggacaag aatcctttcc    420
taggactgct ggacaagagg atcgtagtct aaaaattatc aaaagaacaa agaaacccga    480
cctgcagatc tatcagcctg gacgacgttt gcagactgtt agcaaagaat ccgccagtcg    540
ggtggaggag gaagaagtcc tcaaccaggt agaacaactg agagtagagg aagatgagtg    600
tagggggaaat gttgcgaagg aggaagttgc gaataaacca gacagggccg agatagaaaa    660
gagcccaggt ggtgggagag taggggctgc aaaaggagaa aaggaaaga ggatgggaaa      720
aggggagggg gtgagggaaa cccacgacga cccggcccgc gggaggccgg ctccgcaaa      780
gcgctactcc cgctcagaca aacgaaggaa tcgctaccgc acgcgcagca ccagctcagc     840
```

11

```
tggcagcaac aacagcgctg agggagctgg cctgacggat aatggatgtc gccgccgccg        900
acaggatagg accaaggaga ggccaccact gaagaagcaa gtgtctgtgt cctcaaccga        960
ttccttagac gaggacagaa ttgatgagcc tgatggatta ggacccagga gaagttcaga       1020
aaggaagaga catttagaaa gaaactggtc tggccgtggg gagggtgagc agaaaaccag       1080
tgctaaagaa tatcgaggca ctcttcgtgt cactttcgat gcagaagcca tgaacaaaga       1140
gtctcccatg gtgaggtcag ccagggatga tatggataga ggaaagcctg acaaaggctt       1200
gagcagtggg ggcaaaggct ctgagaagca ggagtccaaa aacccgaaac aagaacttcg       1260
gggtcgtggt cgtggcattc tgattttgcc tgcccatacc accctatctg tcaattcagc       1320
aggttctcca gagtccgcgc ctttgggacc tcggcttttg tttggatctg gtagtaaggg       1380
atctcggagt tggggccgtg gaggcaccac acgccgattg tgggacccaa acaatcctga       1440
tcagaaacct gctctaaaga ctcagacgcc ccagctacat ttcttggaca ctgatgatga       1500
agtcagccct acatcttggg gtgactcacg ccaggctcag gcatcttact ataagtttca       1560
aaactctgac aacccctatt attacccccg gacaccaggc cctgcctccc agtatcccta       1620
tacgggctat aaccctctac agtacccagt gggccctacg aatggtgtgt acccagggcc       1680
ttactaccca ggctacccga ctccgtcagg acagtatgtg tgtagccctc tacctaccag       1740
caccatgagt cccgaggagg tagagcagca catgaggaac ctgcagcaac aggagctgca       1800
caggcttctc cgggtggctg acaaccagga actgcagctc agcaacctgc tctccaggga       1860
ccgcatcagt ccggagggcc tggagaagat ggcgcaactc agagctgaac tgctgcagct       1920
atatgagcgc tgtattctat tagatattga gttctctgat aatcagaatg tggatcagat       1980
cctgtggaag aatgctttct atcaggtgat tgagaagttc aggcaacttg tcaaggatcc       2040
gaatgttgag aacccagaac agattcggaa cagacttttg gagctcttgg atgagggtag       2100
tgacttcttt gatagtttgc ttcagaagct gcaggttact tacaagttca aactggaaga       2160
ctacatggat ggtcttgcca ttcgcagcaa gccattacgc aagacagtaa aatatgcctt       2220
gatcagtgcc cagcgatgca tgatatgcca aggagatatt gctaggtacc gggagcaagc       2280
cagtgataca gcgaattatg ggaaagcacg cagttggtac ctgaaggccc agcacattgc       2340
tcccaagaat gggcgcccct ataaccagtt ggctttgctg gcagtgtata cgaggaggaa       2400
gcttgacgct gtctattact atatgcgcag tttagctgcc agcaacceta tcctgactgc       2460
caaggagagt ctcatgagct tgtttgaaga gaccaagcgg aaggcagaac agatggaaaa       2520
gaagcaacat gaggaatttg acctgagccc tgaccagtgg cggaaaggaa agaagtctac       2580
tttccggcat gttggagatg acaccactcg cctggagatc tggattcatc catcccatcc       2640
acggtcttcc cagggcactg agtctgggaa ggattctgag caagagaatg gctgggcag        2700
cctgagtccc agtgatctga caaaaggtt catcctcagt tttctccatg cccatgggaa       2760
gctgtttacc cggattggga tggagacatt ccctgcagtg gctgagaagg tcctcaagga       2820
gttccaggtg ttactgcagc acagcccctc tcccattgga agtacccgca tgctgcagct       2880
tatgaccatc aatatgtttg cagtacacaa ctcccagctg aaagactgct ctcggagga        2940
gtgccgctct gtgatccagg aacaagccgc agctctgggc ttggccatgt tttctctact       3000
ggtccgccgc tgcacctgct tacttaagga gtccgccaaa gctcagctgt cctctcctga       3060
ggaccaggat gaccaagacg acatcaaggt gtcttccttt gtcccggacc tgaaggagct       3120
gctccccagt gtcaaagtct ggtcagattg gatgctcggc tacccggaca cctggaatcc       3180
tcctcccaca tccctggatc tgcccctcgca tgttgctgtg gatgtatggt cgacgctggc       3240
tgatttctgt aacatactga ctgcagtgaa tcagtctgag gtgccactgt acaaggaccc       3300
ggatgatgac ctcacccttc ttatcctgga agaggatcgg cttctctcgg gctttgtccc       3360
cttgctggct gcccctcagg acccctgcta cgtggagaaa acctcggata aggttattgc       3420
agctgactgc aaaaagggtca cagtgctgaa gtattttctg gaagcccttt gtggacaaga       3480
agagcctctg ctggcattca agggtggaaa gtatgtgtca gtggcacccg tcccagacac       3540
catgggaaag gaaatgggaa gccaagaggg aacacgactg gaagatgagg aggaggatgt       3600
ggtgattgaa gactttgagg aagattcaga ggctgaaggc agcggaggcg aggatgacat       3660
caggagctt cgggccaaga agctggctct ggccaggaag atagctgagc agcagcgtcg       3720
ccaggaaaag atccaggctg tcctggagga ccacagtcag atgaggcaga tggagctcga       3780
aatcagacct ttgttcctcg taccagacac caacggcttc attgaccacc tggccagtct       3840
ggcgcggctg ctggagagca ggaagtacat cctggtggtg cccctcatcg tgatcaatga       3900
gctggacggc ctggccaagg ggcaggagac agaccaccgg ctgggggct acgcccgtgt       3960
ggtacaagag aaggcccgca gtccatcga gttcctcgag cagcgattcg agagtcggga       4020
ctcttgcctg cgagccctga ccagccgtgg caatgaactc gaatccatcg ccttccgcag       4080
tgaggacatc actggccagc tgggtaacaa cgatgatctc atcctgtcct gctgcctcca       4140
ctactgcaaa gacaaggcta aggacttcat gcccgccagc aaagaggagc caatccggct       4200
actgcgggag gtggtgctgt tgacggatga ccggaacctg cgtgtgaagg cgctcacaag       4260
gaatgttcct gtacgggaca tcccagcctt cctcacgtgg gcccaggtgg ctgagggag        4320
ccacactggg gcccccccc cccgtggaac cgttcctgaa aggccaccag gcgcccagtg       4380
tagcacggaa gatgcccacg tgcctgagcc accaatccac ccagacaata aaccatcctc       4440
ttccaaccca cgccacggcc atgctgtggg ggacctgctc ctcacagagc ccctcccaag       4500
```

```
gatcgggcgg aagctgctgg gaccctcctg ggctgccagg atttagcagg gaggtggctg    4560
gctacagcaa cagcagctgg gcaagccaga taggccgccc atgctctcag cctttctccc    4620
tcccccgtct cattccaagg ctgagggagg gccttctcgc ctggggacgc agccactttc    4680
tccagtggag acagggcagg ggttcagagt ttccgtcaga tgcagtgaaa tcacagttcc    4740
ctttcatctt cagaacctct gtcgtgaatg tgttcaagag gctttggtta agtcaggaag    4800
aagtgcccag ggtgtgtgtc cccagtctcc ctgaggcctg gactcgccca tgaacccaag    4860
tcggcttcta gacagcatgt ccctaacagc agccctgggc ccccacctct tctaccatcc    4920
accccagact taccacacac ccttcctgct gctcctcttc ctgcccttat caacctgggt    4980
ccctcacact tcgccagttg cgtcccgtg gacagtcatg agtctagagg aaagggggcat    5040
ctggtctcag gcccgtgctc tcgggtggcc tccacctgct cccttctcc tcactggcct    5100
ttctttccgt ctagcctcct cttcaggaaa tgtcctgact ctcctcagct cccccttcac    5160
ccctccttgc ccgcctaccc tccctccaga atagcccctc acccttcttc cccttctagt    5220
tgatcctttt cacctccctg atccccttca tttcttcacc gcggttcctc gtcataggg    5280
ttctcactct gaactttccc tctctactac ccatggcagg aacctagtac aggtctccca    5340
cccagggcct tccacccTcg gctcctgtgc tgggagaaac ttccaggcgt ggacagccca    5400
gcctgaggca ttccagtgct ggggcaccgt cgcctaacct ggtttctagc tttgccctca    5460
ctccccggaa aaactgacac tgacacaggg gccctttcct tgcccccttta gctggtacct    5520
cagtggggag gcttccttac caagaatgag ttcctgaaac ccagggccag agacaaggac    5580
aacttagggg aagacggggt tttcggtgga gccaggggca aatcttaatg ggaccagtgg    5640
gggatacccc agagcccatg gcctgactgc acagcctgcc tggaggatgg gtgcgcagct    5700
ctgccctccc tgaggcccag gactatgcca gaagcgatgg ggtaccgtgt aggggagcca    5760
aggccagtag tttggggggta ggagtccct agagtctcag aagactgggc tctttggagt    5820
acagggtccc cggcctctcc tttaagattc tctccccagc tggaaggccg atgactgggt    5880
ggtcgggagg gagacccagc tctcctttct gtcccgtttg cagcactggt tttgtttcct    5940
taataaattt ttagttatga aacat                                          5965
```